



THE UNIVERSITY *of* EDINBURGH

## Edinburgh Research Explorer

### The cognitive prerequisites for language

**Citation for published version:**

Smith, K 2018, 'The cognitive prerequisites for language: Insights from iterated learning', *Current Opinion in Behavioral Sciences*, vol. 21, pp. 154-160. <https://doi.org/10.1016/j.cobeha.2018.05.003>

**Digital Object Identifier (DOI):**

[10.1016/j.cobeha.2018.05.003](https://doi.org/10.1016/j.cobeha.2018.05.003)

**Link:**

[Link to publication record in Edinburgh Research Explorer](#)

**Document Version:**

Peer reviewed version

**Published In:**

Current Opinion in Behavioral Sciences

**General rights**

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

**Take down policy**

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact [openaccess@ed.ac.uk](mailto:openaccess@ed.ac.uk) providing details, and we will remove access to the work immediately and investigate your claim.



# The cognitive prerequisites for language: insights from iterated learning

Kenny Smith<sup>a</sup>

<sup>a</sup>*Centre for Language Evolution, School of Philosophy, Psychology and Language Sciences, University of Edinburgh, Edinburgh, UK*

---

## Abstract

Human languages are transmitted by iterated learning: we learn the language of our speech community by observing language use in communicative interaction, and then in turn we produce linguistic behaviours which become the basis for learning in others. Computational and experimental models of iterated learning show that linguistic structure (including compositional structure, which underpins the open-ended expressivity of human language) evolves on a cultural timescale as a result of this iterated learning process. I consider the implications of this work for our understanding of the cognitive capacities required to support linguistic structure, highlighting the importance of the capacities to acquire compositionally-structured meaning-signal mappings from data, and to reason about the minds of others during learning and use.

---

## Structure gives language open-ended expressive power

Uniquely among the communication systems of the natural world, human language allows the open-ended transmission of information: any idea I am capable of entertaining in my mind can be encoded in a linguistic signal and transmitted to your mind, provided we share a common language. Language achieves this open-ended expressivity by combining two features seen separately in other naturally-occurring communication systems [1]. Firstly, language exhibits *semanticity*: we use words and sentences to refer to objects or states of affairs in the world. Second, language is *combinatorial*, at multiple levels — we combine and recombine speech-sounds to form morphemes, and combine and recombine morphemes to build complex words, phrases, and sentences (e.g. the sentence *she jumped* includes 7 phonemes [transcribed ʃ, i, ɔ̃, ʌ, m, p, t], 3 morphemes [*she*, *jump*, and the past-tense morpheme *-ed*], and 2 words).

While these component features are seen elsewhere in the natural world (see Figure 1), human language is unusual in exploiting the combinatorial structure of signals to convey complex meanings. All human languages are pervasively *compositional*: the meaning of a complex signal is a function of the meaning of its parts

---

\*Corresponding author: kenny.smith@ed.ac.uk

and the way in which they are combined [2]. Compositionality allows us to convey differences in meaning by choosing different morphemes to occupy a particular structural position (e.g. *she jumped* means something different from *he jumps* by virtue of the differences in meaning of the elements *she* and *he*, *-ed* and *-s*), or by combining morphemes in different structural configurations (e.g. *Sam annoyed Jess* means something different from *Jess annoyed Sam*, and the meaning of an ambiguous sentence like *she saw the man with the telescope* depends on the structure one assigns to it). All human languages provide a grammar, a system for combining meaning-bearing units in a rule-governed way. Knowing the grammar of a language allows you to encode your thoughts, and (together with the context in which an utterance is produced) decode the encoded thoughts of others; in contrast, as discussed in the caption of Figure 1, the meaning-bearing potential of combinatoriality is hardly exploited in the communication systems of other animals.

### Structure emerges from learning and use

How did our species end up with this unusual and perhaps unique system of communication? Human languages are, like many other human behaviours, culturally transmitted — we learn the language of our speech community by observing language use in communicative interaction, and then in turn we produce linguistic behaviours which become the basis for learning in others. Because they are transmitted through this repeated cycle of learning and use, we should expect languages to evolve to reflect pressures inherent in language learning and linguistic communication: linguistic variants which are easy to acquire and useful for communication should appear and proliferate, while those that are hard to learn or which don't serve people's recurring communicative needs will tend to be replaced by better alternatives.

A growing body of modelling and experimental work demonstrates how this cultural evolutionary perspective can explain fundamental structural properties of natural language as a consequence of biases in learning and use. This work has its roots in computational models of learning and use in simulated populations (cf. the seminal works in refs [15, 16, 17, 18, 19, 20]). More recently, experimental *iterated learning* techniques have been developed to study the same processes in the lab: human participants are trained on miniature languages, then reproduce those languages in a recall task or use them to communicate with another participant (see Figure 2), thereby produce the input for learning by subsequent individuals (see Figure 3). Passing languages along chains of transmission provides a laboratory analogue for the transmission of real languages in the wild, and allows us to identify the conditions under which language-relevant features develop. These experimental techniques have been applied to study the emergence of symbols [21, 22, 23], combinatoriality [24, 25, 26], and compositionality [27, 28, 29]. Here, I focus on work on the evolution of compositional structure, which suggests that compositionality structure emerges when pressures from learning and use are both at play, but not when these pressures operate in isolation.

Languages which are passed from person to person via learning and recall (where the aim of the participant is simply to reproduce the language they were trained on

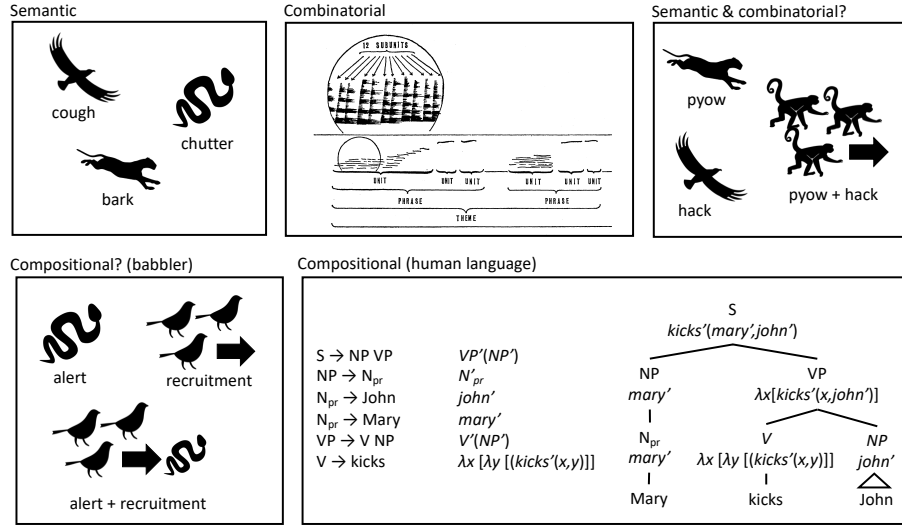


Figure 1: Illustration of cases of semanticity, combinatoriality and compositionality in naturally-occurring communication systems. Semanticity is widespread in the animal kingdom, for example in alarm or food calls [see e.g. 3, 4, for review]; here, semanticity is exhibited by part of the alarm call system of vervet monkeys, who use three distinct alarm calls (glossed “cough”, “chutter” and “bark”) to refer to three predators (eagles, snakes and leopards respectively) [5]. Combinatoriality is also seen in animal communication, in cases where complex vocalisations are built by combining elements from a smaller inventory of recombinable units, for example in birdsong, the long calls of the gibbon, or (as illustrated here) the song of humpback whales [6, 7, 8]. Combinatoriality in these song systems is (as far as we currently know) essentially ornamental and serves no referential function. There is also some evidence of rudimentary combinatoriality and perhaps compositionality in alarm calls. Putty-nosed monkeys combine two distinct alarm call types in sequence to provide a third communicative signal which serves a distinct communicative function (coordinating group movement rather than signalling the presence of predators, [9]), a semantic and combinatorial system which appears also to be exploited to increase the available repertoire of distinct signals in other forest monkeys [10] and even in bacteria [11]. There is also evidence from monkeys and birds that call combinations can have semantic effects beyond merely providing a distinct vocalisation. Southern pied babblers produce two-call combinations consisting of an alert call and a recruitment call, where the combination serves to recruit group members to mob predators [12]. A similar use of two-call combinations is seen in Campbell’s monkeys [13]. While it remains to be seen whether more sophisticated combinatorial and compositional systems will be discovered in other species, current cases therefore clearly lack the open-ended expressivity of human language, which is illustrated here with a toy grammar where each syntactic operation (rules with arrows) is associated with a semantic operation (semantic primitives in *italics*). This grammar allows complex signals with predictable meanings to be constructed and interpreted by recursively applying rules and their associated semantics, producing tree structures as seen on the right where each intermediate construction has a well-defined semantics. Animal icons from thenounproject.com (eagle created by Tatiana Belkina, snake by Jennifer Cozzette, leopard by Adriano Emerick, monkey by Bailey Thompson, bird by Alena Artemova); whale song image from R. S. Payne, S. McVay, Songs of humpback whales, Science 173 (1971) 585– 597, reprinted with permission from AAAS; semantic formalism for human grammar inspired by [14].

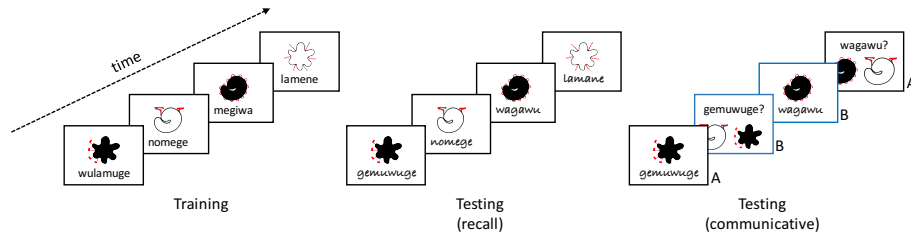


Figure 2: Illustration of components of an artificial language learning paradigm. During training, participants are familiarised with labels for stimuli (here, images of novel shapes, from [28]). Two methods can then be used to extract data from participants for subsequent transmission to a new learner. In a recall test (middle) the participant is simply prompted with stimuli and asked to recall their labels (participant’s productions indicated by handwriting font); errors during recall will result in a new version of the language. In a communicative test, pairs of participants (here labelled A and B) interact communicatively: participant A generates a label for a given stimulus, as in a recall test, which their partner B then has to interpret, e.g. by selecting from an array of possible stimuli; the roles then switch, with B producing a label which is interpreted by A.

to the best of their ability, rather than use it to communicate with another person) tend to become underspecified and eventually highly *degenerate*: multiple related concepts are expressed by a small number of highly ambiguous words [27, 30] (see Figure 3). Learning is the only pressure at play in these experiments, and simpler systems are easier to learn; furthermore, the mistakes learners make tend to increase simplicity (e.g. in reusing a word to convey a set of related meanings, or jettisoning words altogether). The cumulative effect of this simplicity bias in learning is maximally simple languages.

In contrast, languages which are not passed on to naive individuals, but instead used repeatedly for communication by the same pair of individuals, become well-tuned to their communicative needs, but not necessarily compositionally structured ([28]; but see [31] for conditions under which interaction alone can produce structure). The need to communicate successfully forces signals to remain distinct, and since the language is never transmitted to new learners there is no countervailing pressure from learning favouring simplicity.

However, when languages are both learned and used (with the language produced during communication by one pair forming the input to learning by the next pair in a chain of transmission; see Figure 3), languages evolve which encode features of meaning according to compositional rules. Compositional systems constitute a trade-off between the partially-competing pressures from learning and communication, in that compositional grammars are relatively simple yet expressive: the regularities they contain can be exploited by learners during acquisition, but they nonetheless allow meaning to be unambiguously encoded and decoded.

### The capacities underpinning the emergence of structure

What cognitive capacities must be in place for linguistic structure to emerge from iterated learning? A review of the modelling literature plus new modelling

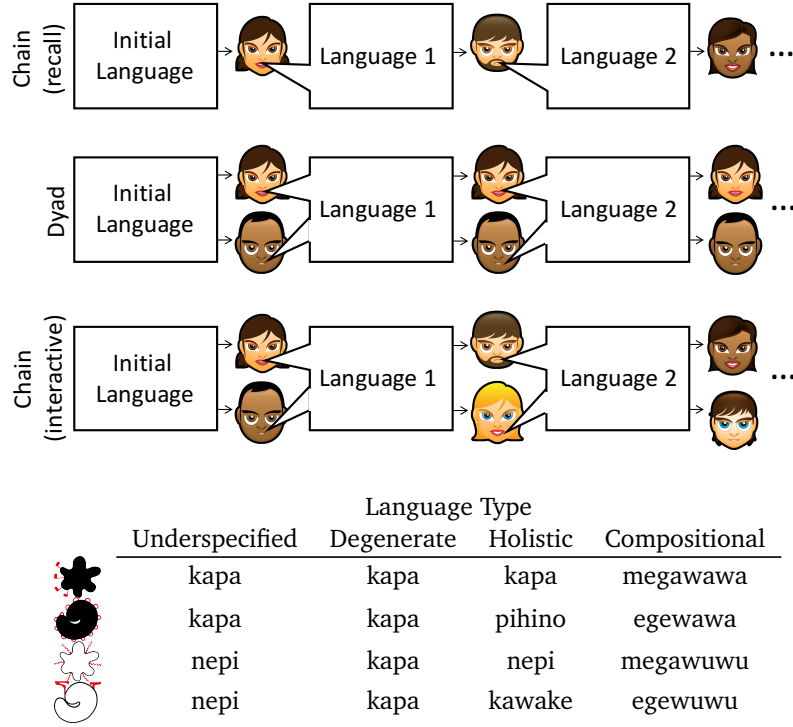


Figure 3: Population configurations for iterated learning (upper) and illustrative languages (lower; note that these example languages are smaller than those used in the cited papers). In recall chains (as in e.g. [27]), the language reproduced by one participant during a recall test becomes the target for learning in another individual. Under these circumstances, underspecified or degenerate language emerge, where the overall number of labels reduces and labels become partially or fully ambiguous. In dyads, a single pair of individuals interact repeatedly; in this condition, holistic languages (which provide a distinct label for each referent but do not exhibit compositional structure) can be preserved. Finally, in interactive chains, the language produced during communication by one pair provides the input to learning in a new pair of individuals; compositional languages emerge in these circumstances (in the example language: first two syllables encode shape, second two syllables encode colour), which allow the language to be relatively easily learned by each new generation but (unlike underspecified or degenerate languages) useful for communication. Stimuli and example languages from [28].

work [28] indicates that there are two main capacities required: the capacity to learn compositional mappings from data, and to reason about the pragmatics of communication.

#### *The capacity to acquire compositional systems*

Unsurprisingly, in order for a compositional system to develop through learning and use, learners have to be able to acquire compositionally-structured systems from input data (although they do not need to be predisposed to do so — in [28] learners were orders of magnitude more likely to infer a degenerate language than a compositional one). The capacity to acquire a compositional language can itself be broken down into three component capacities.

Firstly, it requires the ability to learn the form of signals from input. This capacity is widespread in the natural world, in vocal learners (e.g. song birds, whales, [6, 32]), but also in non-vocal modalities (e.g. in apes trained to use artificial communicative media, [33]). Animal artificial grammar learning studies also show that a range of species are sensitive to regularities in the form of input sequences (e.g. [34, 35]).

Secondly, learning a compositional grammar requires the ability to identify that signals are motivated by or associated with (sometimes complex) internal or environmental stimuli. In humans, this involves identifying the communicative intention behind signals [36, 37], one aspect of our sophisticated capacity to reason about the minds of others [38]. Non-human animals have much more rudimentary capacities for inferring the mental states of others; until recently there was little evidence that even our closest relatives understood that other individuals had mental states which diverged from their own, although recent experimental evidence is overturning this belief [39]. Despite these limitations, various animals have been trained to associate referents with labels: dogs can learn a large inventory of word-object pairings [40, 41], and chimpanzees can be trained to associate visual symbols with objects or colours [42]. It seems unlikely that this involves reasoning about the signaller's communicative intention, and simpler processes of associative or reinforcement learning might suffice in these cases, although presumably inferring more complex communicative intentions requires more sophisticated mindreading abilities.

Finally, learning a compositional language not only involves learning the form of signals and identifying their meaning, but also acquiring and representing a compositional mapping relating signals and meanings, i.e. going beyond non-compositional associations between discrete symbols and atomic concepts to a grammar which specifies how complex combinations can be built and interpreted (as in the compositional grammar in Figure 1). There are cases where non-human animals have been trained on miniature languages which exhibit compositionality. As reviewed in Figure 1, there are alarm call systems arguably featuring this kind of compositionality; however, there is no evidence that the rules of call combination or their semantic effects are themselves learned. Domestic dogs have been trained to respond appropriately to two- and three-word instructions, where doing so requires comprehension of both the component words and the semantics associated with their combination, e.g. understanding the difference between the instructions *take*

*the bone to the ball* and *take the ball to the bone* [43, 44]. Two captive dolphins were trained to respond correctly to more complex, compositional commands [45]; a human-reared chimpanzee was capable of interpreting complex instructions, and also of producing complex signals which exhibited some of the same structures ([33], but see [46] for an intriguing analysis of the limits of Kanzi’s linguistic knowledge); similar abilities have been shown in a language-trained parrot [47].

While these case studies show that non-human animals are capable of learning compositional systems, they shed less light on the process of learning itself: while the procedures for demonstrating the command of the system are rigorously controlled, the training procedures are far less well documented and controlled (naturally so since training spans many years). Some basic questions therefore remain about how these systems are learned by non-humans. What kind of input is required for them to learn compositional grammars? Is mere exposure to complex utterances enough, or is carefully-staged training required, moving from simple one-element sequences to more complex combinations? What kind of errors are made at various stages of learning, and what do those tell us about the biases at play during learning? Much of the animal AGL literature has focussed on the learning of grammars for complex sequences (relevant to the first point above), but studies applying the same rigorous techniques to the acquisition of compositional mappings are virtually non-existent. One exception is a study of the ability of Guinea baboons to learn a simple compositional mapping between coloured shapes and visually-presented two-letter sequences [48], where the colour of the shape was encoded in one letter of the sequence and the shape in the other. The ability of the animals to interpret novel, untrained combinations was surprisingly poor, despite many thousands of training trials; only one of seven baboons showed evidence of generalization, and even then performance seemed rather fragile; furthermore, training on individual letter-colour or letter-shape correspondences prior to training on two-element combinations seemed to hinder acquisition of the compositional mapping, rather than facilitate it. More work is clearly required to build a more comprehensive picture of the conditions under which compositional mappings can be acquired by non-human animals, and whether they bring the same biases to these tasks as humans.

#### *The capacity for pragmatic reasoning*

Once a compositional system has been acquired, it has to be deployed for communication. In our modelling work [28] we assumed that during interaction individuals have a tendency to avoid utterances which were potentially ambiguous; this alone was sufficient to penalise degenerate languages and push the languages towards expressive types. A speaker who models their audience will identify the problems posed by ambiguous utterances and therefore avoid such utterances [49, 37]; humans indeed seem to be capable of substantially more sophisticated inferences about their interlocutor than this [50]. Much less is known about the capacities of non-humans to make similar pragmatic inferences [38, 51], although recent painstaking observational studies of chimpanzee alarm calling in the wild provide suggestive evidence that their use of warning calls is modulated by the audience’s awareness of dangers [52], which would indicate some (rudimentary) capacity to reason about the mental states of others during communication.



### *Human uniqueness in the capacities underpinning the emergence of structure*

Perhaps surprisingly, at least the rudiments of all of the capacities required to support the cultural evolution of compositional, language-like systems are present in non-humans, including in other great apes, suggesting that the discontinuities between humans and other animals in the capacities underpinning language may be rather minimal. Four observations seem worth making. Firstly, the ability to reason about the mental states of others plays a dual role in the evolution of structure: reasoning about the minds of others allows a learner to infer the intended meaning behind an utterance, and then dictates how the inferred language is deployed during communication. Understanding the evolution of the human capacity for theory of mind therefore seems likely to be highly informative to understanding the evolution of language, and language and theory of mind might indeed have co-evolved [53]. Second, the ability to produce complex signals seems unlikely to be the limiting factor, since non-vocal modalities have the same combinatorial potential. Third, the conditions under which non-human animals can acquire compositional mappings seems like a major gap in our current knowledge, and a potentially informative line for future animal AGL work. Fourth, and finally, one prediction of the cultural evolutionary approach is that if the ‘missing’ capacities could be scaffolded in some other species, then we would expect the same types of structure to evolve, a prediction which can be tested experimentally [54].

### **Conclusions**

Understanding the conditions under which structured linguistic systems emerge from iterated learning, the repeated cycle of learning and use, provides useful insights into the cognitive capacities required for language. Our approach has been to model (in simulation or in the lab) the way in which linguistic systems evolve. This approach suggests two lines for future research on non-human cognition, in particular, work which will uncover the cognitive capacities, evolutionary trajectories and selection pressures leading to the capacity to learn compositional grammars and to make inferences during communication about how they should be used.

### **Acknowledgements**

This project has received funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (grant agreement no 681942).

### **Annotations for bibliography**

*For reference [28] (one star)*

Computational and experimental models of iterated learning, instantiating the various contrasts between population types shown in in Figure 3. Demonstrates that linguistic structure emerges when pressures inherent in learning and use are *both* at play, but either in isolation is insufficient.

*For reference [39] (two stars)*

In the main paper I argue that capacity to represent and reason about the mental states of others is a key prerequisite for the evolution of linguistic structure. This paper shows for the first time the ability to reason about false belief in great apes, and therefore represents a major step in our understanding of social cognition in non-humans. To demonstrate false belief understanding, they adapted anticipatory looking paradigms from infant research, and showed that great apes anticipated the actions of another individual based on that individual's (false) beliefs about the world.

*For reference [48] (one star)*

In the main paper I argue that the capacity of non-human animals to acquire compositional mappings from input is an important yet understudied topic. This paper presents two experiments exploring systematically the ability of baboons to acquire a simple compositional mapping between coloured shapes and visually-presented two-letter sequences. The ability of the animals to interpret novel, untrained combinations was surprisingly poor.

*For reference [37] (two stars)*

This is a review of the Rational Speech Act Model, a formalisation of Grice's work on the pragmatics of communication. The Rational Speech Act model provides a formalisation of the inference of meaning in context, and provides an indispensable set of tools for researchers who want to model the process by which linguistic systems are built through interaction and adapt to their context of use.

*For reference [52] (one star)*

In the main paper I argue that capacity to represent and reason about the mental states of others is a key prerequisite for the evolution of linguistic structure. This paper provides suggestive (but not iron-clad) evidence from wild chimpanzees that chimpanzees are capable of reasoning about knowledge and ignorance in their audience, as evidenced by the effects of audience knowledge on their alarm calling behaviour.

## References

- [1] C. F. Hockett, The origin of speech, *Scientific American* 203 (1960) 88–96.
- [2] M. Krifka, Compositionality, in: R. A. Wilson, F. Keil (Eds.), *The MIT Encyclopedia of the Cognitive Sciences*, MIT Press, Cambridge, MA, 2001.
- [3] S. W. Townsend, M. B. Manser, Functionally referential communication in mammals: The past, present and the future, *Ethology* 119 (2013) 1–11.
- [4] S. A. Gill, A. M.-K. Bierema, On the meaning of alarm calls: A review of functional reference in avian alarm calling, *Ethology* 119 (2013) 449–461.

- [5] D. Cheney, R. Seyfarth, *How Monkeys See the World: Inside the Mind of Another Species*, University of Chicago Press, Chicago, IL, 1990.
- [6] C. K. Catchpole, P. J. B. Slater, *Bird Song: Biological Themes and Variations*, Cambridge University Press, Cambridge, 1995.
- [7] J. J. Raemaekers, P. M. Raemaekers, E. H. Haimoff, Loud calls of the gibbon (*Hylobates lar*): repertoire, organization and context, *Behaviour* 91 (1–3) (1984) 146–189.
- [8] R. S. Payne, S. McVay, Songs of humpback whales, *Science* 173 (1971) 585–597.
- [9] K. Arnold, K. Zuberbühler, Meaningful call combinations in a non-human primate, *Current Biology* 18 (2008) R202–R203.
- [10] K. Ouattara, A. Lemasson, K. Zuberbühler, Campbell’s monkeys use affixation to alter call meaning, *PLOS ONE* 4 (2009) e7808.
- [11] T. C. Scott-Phillips, J. Gurney, A. Ivens, S. P. Diggle, R. Popat, Combinatorial communication in bacteria: Implications for the origins of linguistic generativity, *PLOS ONE* 9 (2014) e95929.
- [12] S. Engesser, A. R. Ridley, S. Townsend, Meaningful call combinations and compositional processing in the southern pied babbler, *Proceedings of the National Academy of Sciences, USA* 113 (2016) 5976–5981.
- [13] K. Zuberbühler, A syntactic rule in forest monkey communication, *Animal Behaviour* 63 (2) (2002) 293–299.
- [14] R. Cann, *Formal Semantics: an introduction*, Cambridge University Press, Cambridge, 1993.
- [15] J. R. Hurford, Biological evolution of the saussurean sign as a component of the language acquisition device, *Lingua* 77 (2) (1989) 187–222.
- [16] S. Kirby, Learning, bottlenecks and the evolution of recursive syntax, in: E. Briscoe (Ed.), *Linguistic Evolution through Language Acquisition: Formal and Computational Models*, Cambridge University Press, Cambridge, 2002, pp. 173–203.
- [17] L. Steels, The origins of syntax in visually grounded robotic agents, *Artificial Intelligence* 103 (1998) 133–156.
- [18] B. de Boer, Self-organization in vowel systems, *Journal of Phonetics* 28 (2000) 441–465.
- [19] P.-Y. Oudeyer, The self-organization of combinatoriality and phonotactics in vocalization systems, *Connection Science* 17 (2005) 325–341.

- [20] A. Wedel, Lexical contrast maintenance and the organization of sublexical contrast systems, *Language and Cognition* 4 (2012) 319–355.
- [21] S. Garrod, N. Fay, J. Lee, J. Oberlander, T. MacLeod, Foundations of representation: where might graphical symbol systems come from?, *Cognitive Science* 31 (2007) 961–987.
- [22] C. A. Caldwell, K. Smith, Cultural evolution and perpetuation of arbitrary communicative conventions in experimental microsocieties, *PLOS ONE* 7 (2012) e43807.
- [23] N. Fay, T. M. Ellison, The cultural evolution of human communication systems in different sized populations: Usability trumps learnability, *PLOS ONE* 8 (2013) e71781.
- [24] T. Verhoef, S. Kirby, B. de Boer, Emergence of combinatorial structure and economy through iterated learning with continuous acoustic signals, *Journal of Phonetics* 43 (2014) 57–68.
- [25] G. Roberts, J. Lewandowski, B. Galantucci, How communication changes when we cannot mime the world: Experimental evidence for the effect of iconicity on combinatoriality, *Cognition* 141 (2015) 52–66.
- [26] H. Little, K. Eryilmaz, B. de Boer, Signal dimensionality and the emergence of combinatorial structure, *Cognition* 168.
- [27] S. Kirby, H. Cornish, K. Smith, Cumulative cultural evolution in the laboratory: an experimental approach to the origins of structure in human language, *Proceedings of the National Academy of Sciences, USA* 105 (2008) 10681–10686.
- [28] S. Kirby, M. Tamariz, H. Cornish, K. Smith, Compression and communication in the cultural evolution of linguistic structure, *Cognition* 141 (2015) 87–102.
- [29] C. Beckner, J. B. Pierrehumbert, J. Hay, The emergence of linguistic structure in an online iterated learning task, *Journal of Language Evolution* 2 (2017) 160–176.
- [30] C. Silvey, S. Kirby, K. Smith, Word meanings evolve to selectively preserve distinctions on salient dimensions, *Cognitive Science* 39 (2015) 212–226.
- [31] J. Winters, S. Kirby, K. Smith, Linguistic systems adapt to their contextual niche, *Language and Cognition* 7 (2015) 415–449.
- [32] E. C. Garland, L. Rendell, L. Lamoni, M. M. P. M. J. Noad, Song hybridization events during revolutionary song change provide insights into cultural transmission in humpback whales, *Proceedings of the National Academy of Sciences, USA* 114 (2017) 7822–7829.

- [33] S. Savage-Rumbaugh, K. McDonald, R. A. Sevcik, W. D. Hopkins, E. Rubert, Spontaneous symbol acquisition and communicative use by pygmy chimpanzees (*Pan paniscus*), *Journal of Experimental Psychology: General* 115 (3) (1986) 211–235.
- [34] W. T. Fitch, M. D. Hauser, Computational constraints on syntactic processing in a nonhuman primate, *Science* 303 (2004) 377–380.
- [35] A. E. Milne, C. I. Petkov, B. Wilson, Auditory and visual sequence learning in humans and monkeys using an artificial grammar learning paradigm, *Neuroscience*.
- [36] M. C. Frank, N. D. Goodman, Inferring word meanings by assuming that speakers are informative, *Cognitive Psychology* 75 (2014) 80–96.
- [37] N. D. Goodman, M. C. Frank, Pragmatic language interpretation as probabilistic inference, *Trends in Cognitive Sciences* 20 (2016) 818–829.
- [38] T. C. Scott-Phillips, *Speaking Our Minds*, Palgrave Macmillan, London, 2014.
- [39] C. Krupenye, F. Kano, S. Hirata, J. Call, M. Tomasello, Great apes anticipate that other individuals will act according to false beliefs, *Science* 354 (2016) 110–114.
- [40] J. Kaminski, J. Call, J. Fischer, Word learning in a domestic dog: evidence for “fast mapping”, *Science* 304 (2004) 1682–1683.
- [41] J. W. Pilley, A. K. Reid, Border collie comprehends object names as verbal referents, *Behavioural Processes* 86 (2011) 184–195.
- [42] T. Asano, T. Kojima, T. Matsuzawa, K. Kubota, K. Murofushi, Object and color naming in chimpanzees (*pan troglodytes*), *Proceedings of the Japanese Academy* 58 (1982) 118–122.
- [43] J. W. Pilley, Border collie comprehends sentences containing a prepositional object, verb, and direct object, *Learning and Motivation* 44 (2013) 229–240.
- [44] D. Ramos, C. Ades, Two-item sentence comprehension by a dog (*canis familiaris*), *PLOS ONE* 7 (2012) e29689.
- [45] L. M. Herman, D. G. Richards, J. P. Wolz, Comprehension of sentences by bottlenosed dolphins, *Cognition* 16 (129–219).
- [46] R. Truswell, Dendrophobia in bonobo comprehension of spoken english, *Mind and Language* 32 (2017) 395–415.
- [47] I. M. Pepperberg, Functional vocalizations by an african grey parrot (*psittacus erithacus*), *Zeitschrift für Tierpsychologie* 55 (1981) 136–160.
- [48] T. Medam, J. Fagot, Behavioral assessment of combinatorial semantics in baboons (*papio papio*), *Behavioural Processes* 123 (2016) 54–62.

- [49] M. C. Frank, N. D. Goodman, Predicting pragmatic reasoning in language games., *Science* 336 (2012) 998.
- [50] H. P. Grice, Logic and conversation, in: P. Cole (Ed.), *Syntax and Semantics*, Vol. 3, Academic Press, New York, 1975, pp. 41–58.
- [51] T. C. Scott-Phillips, Pragmatics and the aims of language evolution, *Psychonomic Bulletin and Review* 24 (2017) 186–189.
- [52] A. M. Schel, S. W. Townsend, Z. Machanda, K. Zuberbühler, K. E. Slocumbe, Chimpanzee alarm call production meets key criteria for intentionality, *PLOS ONE* 8 (2013) e76674.
- [53] M. Woensdregt, K. Smith, Pragmatics and language evolution, in: *Oxford Research Encyclopedia of Linguistics*, Oxford University Press, 2017.
- [54] N. Claidière, K. Smith, S. Kirby, J. Fagot, Cultural evolution of a systematically structured behaviour in a non-human primate, *Proceedings of the Royal Society of London B* 281 (2014) 20141541.